Tests of the Expectations Hypothesis: Resolving the Campbell-Shiller Paradox

Authors	Daniel L. Thornton
Working Paper Number	2003-022B
Revision Date	
Citable Link	https://doi.org/10.20955/wp.2003.022
Suggested Citation	Thornton, D.L., 2003; Tests of the Expectations Hypothesis: Resolving the Campbell-Shiller Paradox, Federal Reserve Bank of St. Louis Working Paper 2003-022. URL https://doi.org/10.20955/wp.2003.022

Published In	Journal of Money, Credit, and Banking

Tests of the Expectations Hypothesis: Resolving the Campbell-Shiller Paradox

Daniel L. Thornton
Federal Reserve Bank of St. Louis
Phone (314) 444-8582
FAX (314) 444-8731
Internet Address: thornton@stls.frb.org

August 2004

Abstract

One of the more puzzling results in the expectations hypothesis (EH) testing literature is the Campbell-Shiller paradox. In an influential paper, Campbell and Shiller (1991) found that "the slope of the term structure almost always gives a forecast in the wrong direction for the short-term change in the yield on the longer bond, but gives a forecast in the right direction for long-term changes in short rates." This paper provides an econometric resolution to the Campbell-Shiller paradox. Specifically, it shows that, by their construction, these tests can generate results consistent with the Campbell-Shiller paradox if the EH does not hold—whatever the reason. Monte Carlo experiments confirm that this explanation can account for Campbell and Shiller's paradoxical results for most pairings of short-term and long-term rates considered.

JEL Classification: E40, E52
Key Words: expectations hypothesis, Campbell-Shiller Paradox, Monte Carlo experiment

The views expressed here are the author's and do not necessarily reflect the views of the Board of Governors of the Federal Reserve System or the Federal Reserve Bank of St. Louis. I would like to thank John Campbell, Martin Sola, an anonymous referee, and the editor, Masao Ogaki, for helpful comments and Jonathan Ahlbrecht, Stephen Majesky, William Bock, Kent Koch, Stephen Gardner, and John Zhu for valuable research assistance.

1. Introduction

The expectations hypothesis (EH) of the term structure of interest rates is the proposition that the long-term rate is determined by the market's expectation for the short-term rate over the holding period of the long-term asset, plus a constant risk premium. The EH, which asserts that expected excess returns are time invariant, plays an important role in economics and finance, especially in monetary policy analyses. Hence, it is not surprising that the EH has been painstakingly investigated. One of the more troubling results in the literature is the Campbell-Shiller paradox (CSP). In their influential work, Campbell and Shiller (1991) found that, while the EH was nearly always rejected, the results differed markedly depending on the test used. One of these tests, which I call the *conventional test* because of its frequent use, regresses the long-term change in the short-term rate on the spread between the long-term and short-term rates. The EH is tested by testing the hypothesis that the slope coefficient, β, equals 1—its theoretical value if the EH holds. While the EH is frequently rejected, estimates of β are typically positive and significantly different from zero. Moreover, estimates of $\overline{R}^2$ are frequently significantly different from zero, suggesting that the slope of the yield curve has predictive power for the short-term rate.

The other test, which I call the *contrarian test* because of its well-documented tendency to generate results contrary to the EH, regresses the short-term change in the long-term rate on the rate spread. Estimates of the slope coefficient, λ, (which should also equal 1 if the EH holds) are frequently negative, and estimates of $\overline{R}^2$ are often insignificantly different from zero. Hence, the rate spread not only suggests the "wrong" direction for the short-term change in the long-term rate, but it explains virtually none of it.

Campbell and Shiller (1991, p. 505) concluded that

> We thus see an apparent paradox: the slope of the term structure almost always gives a forecast in the wrong direction for the short-term change in the yield on the longer bond, but gives a forecast in the right direction for long-term changes in short rates.

There have been a number of attempts to generate models of the term structure that are consistent with the observation that the slope of the term structure almost always gives forecasts the wrong direction for the short-term change in the long-term rate, but in the right direction for the long-term change in the short-term rate by attributing the failure of the EH to a specific cause—e.g., the EH holds but, (i) the risk premium is time-varying, (ii) expectations are not rational, (iii) the long-term rate over-reacts to expected changes in the short-term rate, etc., (e.g., Fama, 1984; Mankiw and Miron, 1986; Backus, et al., 1989; Froot, 1989; Simon, 1990; Campbell and Shiller, 1991; Hardouvelis, 1994; McCallum, 1994a; Campbell, 1995; Dotsey and Otrok, 1995; Roberds et al., 1996; Balduzzi, et al. 1997; Bekaert et al., 1997, 2001; Hsu and Kugler, 1997; Tzavalis and Wickens, 1997; Driffill et al., 1997; Roberds and Whiteman, 1999; Bansal and Zhou, 2002; and Dai and Singleton, 2002).

The desire to attribute the CSP to a specific failure of the EH is understandable. Clearly, long-term financial markets must be forward looking, so there must be some sense in which the long-term rate is determined by the market's expectation for the future level of rates. Ideally, the rejection of the EH would be attributed to a single cause. Nevertheless, in its desire to account for the CSP, the literature has overlooked a relatively simple, econometric resolution of the CSP—namely, that these tests are capable of generating results that are consistent with the CSP when the EH does not hold *whatever the reason*. This agnostic approach answers the question: Why does the conventional test generate estimates that appear favorable to the EH, while the contrarian test (applied to the same data) generates results that are completely at odds with it? This is the essence of the CSP. Moreover, because the EH is nearly always rejected, it resolves the CSP is a manner that is consistent with the literature itself. Unfortunately, the broader question of the why EH does not hold remains unresolved.

Knowing that these tests generate results that are consistent with the CSP when the EH does not hold has implications for applied work. First, knowing that the estimate of β is no more reliable an indicator of the validity of the EH than the estimate of λ is evidence against it, researchers will be careful not to use the magnitude of the estimate of β (or the corresponding estimate of $\overline{R}^2$) as an indicator of the validity of the EH, as has been done (e.g., Mankiw and Miron, 1986; Campbell and Shiller, 1991; Roberds et al., 1996; Balduzzi et al., 1997; Tzavalis and Wickens, 1997; Kozicki and Tinsley, 2001). This implication applies to tests of other financial theories based on similar equations (e.g., Campbell and Shiller, 1987; McCallum, 1994b; Baillie and Bollerslev, 2000).[1]

Second, these results will hopefully focus attention away from using these single equation tests of the EH and stimulate the use of alternative tests of the EH, such as the multivariate test proposed by Campbell and Shiller (1987), which has been made operational recently by Bekaert and Hodrick (2001). This methodology tests the restrictions implied by the EH on a general vector autoregression (VAR) representation of the short-term and long-term rates. This test is more difficult to employ than the single equation tests that generate the CSP; however, because the VAR encompasses a wider array of alternative hypotheses than single equation tests derived under the null, it should

[1] Because of its tendency to predict the wrong direction for the short-term change in the long-term rate, researchers never based their conclusions about the EH on the estimate of λ.

be more powerful. Moreover, the test can be applied to a VAR with more than two interest rates or that includes other economic variables which might condition market expectations. It can also test the EH under the assumptions that interest rates are either stationary or non-stationary.

Third, knowing that by their construction these tests frequently generate the CSP results when the EH does not hold for any reason will refocus the attention away from explaining the CSP to explaining why the EH does not hold. Hence, rather than positing a particular failure of the EH and investigating whether this failure can account for the CSP, as much of the literature has done, the focus will hopefully shift to alternative ways to investigate why the EH does not hold. For example, since the predictability of the "short-term interest rate" is the essence of the EH, researchers might investigate why interest rates are difficult to predict (e.g., Diebold and Li, 2003; and Duffee, 2002).

The resolution of the CSP presented here is motivated by the observation that estimates of β and λ need not be zero when the EH does not hold. Indeed, the estimate of β is positively biased—biased above zero—when the EH does not hold, while the estimate of λ is negatively biased. While the analysis presented here has implications for the power of these tests, the power of these tests is not the issue. Indeed, as we will see, these tests have relatively high power, at least in the cases covered by the experiment done here.

The extent to which these features of these tests account for the CSP is investigated by a Monte Carlo experiment. Specifically, a VAR for all combinations of short-term and long-term rates used by Campbell and Shiller (1991) is estimated by imposing restrictions that guarantee that the EH does not hold. Hypothetical long- and short-term interest rates are generated from this VAR and the conventional and contrarian tests are applied to these hypothetical data. The results obtained using the actual data are then compared with the distributions of the estimates of β and λ obtained from 10,000 replications of the model. The results indicated that, for most combinations of long-term and short-term rates, estimates of β and λ obtained from the historical data could have been obtained by applying these tests to data for which the EH does not hold.

The outline of the paper is as follows. Section 2 derives the conventional and contrarian tests and demonstrates the CSP by applying these tests to an updated sample of the data used by Campbell and Shiller (1991). Previous resolutions of the CSP are discussed briefly in Section 3. Section 4 presents an alternative resolution of the CSP. Section 5 investigates the extent to which this resolution accounts for the results presented in Section 2. The conclusions and implications are presented in Section 6.

2. *The Conventional and Contrarian Tests of the EH*

The EH of the term structure is a proposition about the relationship between a long-term, n-period interest rate, R_t^n, and expected future levels of a short-term, m-period rate, R_t^m, n-m = (k-1)m periods in the future, where $k = n/m$ is an integer. That is,

$$R_t^n = (1/k)\sum_{i=0}^{k-1} E_t R_{t+mi}^m + \pi^{n,m}. \tag{1}$$

Equation 1 states that the n-period rate is equal to the average of the market's expectation for the m-period rate over the term of the n-period rate. The constant risk premium, $\pi^{n,m}$, may vary with the maturity of the long-term and short-term rates.[2]

Both the conventional and contrarian tests are derived under the assumptions that the EH is the true data generating process (DGP) and expectations are rational, i.e.,

$$E_t R^m_{t+mi} = R^m_{t+mi} + \nu_{t+mi}, \quad i = 0,1,...,k-1, \tag{2}$$

where ν_{t+mi} is a mean-zero, *iid* white noise error. The conventional test is derived by substituting (2) into (1) to yield

$$R^n_t = (1/k)\sum_{i=0}^{k-1} R^m_{t+mi} + (1/k)\sum_{i=0}^{k-1} \nu_{t+mi} + \pi^{n,m}. \tag{3}$$

The EH could be tested by rearranging (3) and paramertizing, i.e.,

$$(1/k)\sum_{i=0}^{k-1} R^m_{t+mi} - R^n_t = \alpha + \beta' R^n_t - (1/k)\sum_{i=0}^{k-1} \nu_{t+mi}, \tag{4}$$

and testing the hypothesis that $\beta' = 0$. This is seldom done, however, because interest rates are unit-root or, perhaps more correctly, near-unit-root processes.[3] Rather, the conventional test is obtained by subtracting the short-term rate from both sides (3), which yields

$$(1/k)\sum_{i=0}^{k-1} R^m_{t+mi} - R^m_t = -\pi^{n,m} + (R^n_t - R^m_t) + \omega_t \,.^{4} \tag{5}$$

The conventional test of the expectations theory is then obtained by *parameterizing* (5), which yields

$$(1/k)\sum_{i=0}^{k-1} R^m_{t+mi} - R^m_t = \alpha + \beta(R^n_t - R^m_t) + \omega_t. \tag{6}$$

The EH is tested by estimating (6) and testing the hypothesis $\beta = 1$.

To see the implications of the EH for the expected short-term change in the long-term rate, note that (1) can be rewritten as

$$(n-m)E_t R^{n-m}_{t+m} = nR^n_t + mR^m_t + n\pi^{n,m}. \tag{7}$$

[2] Shiller et al., (1983) argue that Equation 1 is exact in some special cases and that it can be derived as a linear approximation to a number of nonlinear expectations theories of the term structure.

Subtracting $(n-m)R_t^n$ from both sides of (7) and a little algebra yields

$$E_t R_{t+m}^{n-m} - R_t^n = (m/(n-m))(R_t^n - R_t^m) + (n/(n-m))\pi^{n,m}. \tag{8}$$

Again, assuming (2) and parameterizing the resulting expression yields the contrarian test of the EH, i.e.,

$$R_{t+m}^{n-m} - R_t^n = \mu + \lambda(m/(n-m))(R_t^n - R_t^m) + \eta_t. \tag{9}$$

The EH is tested by estimating (9) and testing the hypothesis $\lambda = 1$.

2.1 *The CSP*

Campbell and Shiller (1991) obtained their paradoxical results by estimating (6) and (9) using continuously compounded yields on riskless pure discount bonds, calculated by McCulloch (1990) over the period 1952.01 to 1987.02.[4] The maturity of the bonds range form 1 to 120 months and the tests were applied to all combinations of rates where k is an integer. The CSP is replicated here using an update of these data (McCulloch and Kwon, 1993) for the period 1952.01 to 1991.02.

The results for the conventional and contrarian tests are reported in Tables 1 and 2, respectively. Each cell of the tables reports the estimate of the slope coefficient, the estimate of $\overline{R}^2$, and the significance levels for Wald test of the null hypotheses that the slope coefficients equal 0 and 1, respectively.[5] The results are very similar to those reported by Campbell and Shiller (1991) for the sample period 1952.01 to 1987.02.

---

[3] An exception is Longstaff (2000).

[4] Similar results are obtained for a variety of countries, data sets and time periods (e.g., Fama, 1984; Mankiw and Miron, 1986; Hardouvelis, 1988 and 1994; Simon, 1990; Campbell, 1995; Roberds, et al., 1996; Bekaert et al., 1997; Roberds and Whiteman, 1999).

[5] The standard errors are based on the Hansen-Hodrick (1980) method. However, there were three instances where the estimated variances were negative. In these three instances, the Newey-West (1987) method was used. It is well known that the Hansen-Hodrick and Newey-West procedures can generate misleading results when the degree of overlap is large relative to the sample size, and the Wald test may be poorly sized. These statistics are used here because they are commonly reported statistics in applied work.

In all but one instance, the estimate of β is positive and, in the majority of cases, estimates of β are significantly different from zero. Moreover, as was the case with Campbell and Shiller (1991), estimates of β are greater than 1 when the long-term rate is the ten-year rate and the EH is easily rejected at the shorter end of the term structure, but less frequently at the longer end. For a given maturity of the short-term rate, estimates of β are larger at the short and long ends of the maturity spectrum and smaller in the intermediate range. This is what Campbell and Shiller and others have characterized as a "U-shape" pattern in the estimates of β or what Roberds and Whiteman (1999) term the "smile." This pattern is illustrated in Figure 1, which shows the estimate of β for each long-term rate for $m = 1$. Such results were interpreted as suggesting that the EH works better at the short and long ends of the maturity spectrum and less well in the intermediate maturity range (e.g., Campbell and Shiller, 1991, and Campbell, 1995).

Estimates of λ, presented in Table 2, are again similar to those reported by Campbell and Shiller (1991). There is a preponderance of "wrong signs" in estimates of λ. Indeed, estimates of λ are negative for every combination of long-term and short-term rates save two—0.003 and 3.581 when (n,m) is $(2,1)$ and $(120,60)$, respectively.[6] In addition, with the exceptions of $n = 120$, the estimates of $\overline{R}^2$ suggest that the rate spread explains virtually none of the short-term change in the long-term rate. For nearly every rate pair, the hypothesis $\lambda = 1$ is rejected at a low significance level. Moreover, for a given value of m, estimates of λ tend to become more negative as n increases. Roberds and Whiteman (1999) refer to this pattern as the "smirk." The smirk is illustrated in Figure 2, which shows the estimate of λ for each long-term rate for $m = 1$.

[6] Campbell and Shiller's estimates of λ for the shorter sample were 0.002 and 4.575.

3. Previous Resolutions of the CSP

There have been numerous attempts to attribute one or both of these paradoxical results to a specific violation of the EH. Perhaps the most common candidate is a time-varying risk premium. It is easy to show that if the risk premium is time varying (rather than constant as the EH hypothesizes) and positively correlated with the slope of the yield curve, estimates of β and λ will both be biased downward from their theoretical value of 1. Hardouvelis (1994), however, argues that the time-varying risk premium is implausible and inconsistent with his evidence on its relative variability. He concludes that the time-varying risk premium explanation cannot account for the negative estimates of λ reported for U.S. data. He suggests instead that the "overreaction hypothesis" of Campbell and Shiller (1991) and Froot (1989) is a more likely explanation.

The overreaction hypothesis asserts that the long-term rate overreacts to expected changes in the short-term rate. For example, when some event causes the market to raise its expectation for the future short-term rate, the long-term rate overreacts causing the current spread between the long-term and short-term rate to increase by an amount that is larger than is warranted by future changes in the short-term rate. Hence, over time, the long-term rate falls as short-term rate rises. This creates negative correlation between both short-term changes in the long-term rate and long-run changes in the short-term rate and the yield spread. The negative bias in estimates of λ is magnified by the fact that the yield spread in (9) is weighted by $(m/(n-m))$, so the bias is larger, the larger the difference in maturity between the long-term and short-term rates.

The overreaction hypothesis is supported by Froot (1989), who used survey data and found that the negative correlation between short-term changes in the long-term rate and the yield spread are not the result of a time-varying risk premium, but are due to a violation of the rational expectations assumption. Hardouvelis (1994) also finds that the overreaction hypothesis is a better explanation of the contrarian test anomalies than several alternatives, including a time-varying risk premium.

Bekaert et al. (2001) investigated the possibility that the term structure anomalies are the result of a rational version of the overreaction hypothesis. Specifically, they investigate whether these anomalies are the result of a "peso problem," where high-interest-rate regimes occur less frequently than rationally anticipated. Combining their peso problem with a time-varying risk premium they are able to generate results that are "more consistent with the data." They conclude, however, that their model "cannot fully account for U.S. term structure anomalies."[7]

Balduzzi et al. (1997) suggest that the rejection of the EH is due to the market's inability to predict monetary-policy-induced changes in short-term rates. They find that accounting for this possibility results in a larger estimate of β, but the EH is rejected. No attempt was made to account for the negative estimates of λ.

Campbell (1995) suggests that the CSP anomalies can be explained by "changing rational expectations about excess long bond returns." When the yield spread is "high," market participants must either expect the yield on long-term bonds to be higher over its life than that of a series of short-term bonds, or they must expect short-term rates to rise. The EH rules out the former, i.e., the yield spread only reflects rationally anticipated changes in short-term rates. Hence, Campbell (1995) argues that changing rational expectations of excess long-bond returns acts as a measurement error (e.g., Hardouvelis, 1994; Stambaugh, 1988), biasing the estimates of both β and λ. Because changing rational expectations of excess long-bond returns affects both sides of (9)—positively on the rhs and negatively on the lhs—the downward bias in the estimate of λ may be severe, causing the coefficient to be negative.

In contrast, the changing rational expectations of excess long-bond returns only affects the rhs of (6); hence, the bias will be less severe depending, of course, on the correlation between excess long-bond returns and long-term changes in the short-term rate. Indeed, if this correlation were negative, the estimates of β would be biased upward. Campbell suggests that the combination of predictable policy actions at short horizons (e.g., Balduzzi et al., 1997), policymakers' incentive to make short-term rates predictable at long horizons (Goodfriend, 1991, Rudebusch, 1995 and Woodford, 2001), and interest rate smoothing in the "medium run" may account for the smile illustrated in Figure 1.

The interest rate smoothing explanation for the failure of the EH using (3) was first suggested by Mankiw and Miron (1986). They argued that the Fed's penchant for smoothing short-term interest rates accounts for both the well-documented decline in the predictability of changes in short-term interest rates and the smaller estimate of β shortly after the Fed's founding.

There have been several attempts to account for the CSP anomalies using affine models of the term structure. Roberds and Whiteman (1999) worked out the implications of a one-factor model of the term structure developed by Cox et al. (1985) and Chen and Scott (1992, 1993) for the ability of the yield spread to predict future interest rates. They show that explaining the CSP in this model requires the ratio of the standard deviation of the term premium to the standard deviation of forecast changes in the short or long rates to be only slightly larger then unity—considerably smaller than Hardouvelis' (1994) estimate

[7] Bekaert et al. (2001, p. 241).

of 31.7 for the United States. Nevertheless, they found that certain parameterizations of a two-factor model were broadly consistent with the CSP. Unfortunately, the parameter estimates were inconsistent with maximum likelihood estimates reported by Chen and Scott (1993) and the hypothetical data generated from these models were inconsistent with many of the features of the historical data. Their findings are consistent with those of Backus et al. (1989) and Bansal and Zhou (2002). Bansal and Zhou conclude that such models, including affine specifications with up to three factors, "are sharply rejected by the data." They found, however, that their two-factor regime shift model—with a regime-dependent risk premium—could account for negative estimates of λ, though not the smirk *per se*.[8]

In a similar vein, Dai and Singleton (2002) and Tzavalis and Wickens (1997) show that elements of the CSP can be resolved by allowing sufficient flexibility in capturing the time-varying risk. Dai and Singleton accomplish this by showing that the typical estimates of λ can be explained by a large subclass of dynamic term structure models that allow for flexibility in the specification of the market price of risk. Tzavalis and Wickens, on the other hand, assume that the term premiums associated with different maturities are determined by a common factor. Both Dai and Singleton and Tzavalis and Wickens show that when their estimate of the risk premium is included in (9), the estimate of λ is insignificantly different from unity.[9]

With the exception of these statistical models that are flexible enough to capture much of the observed variation in the term premiums, the literature has not been successful

[8] While their estimates of λ were negative for all of the combinations of long-term and short-term rates considered, their estimates were frequently very different from the estimates obtained using the historical data.

in explaining the CSP. Indeed, Bekaert and Hodrick (2001) note that “the literature has had surprisingly little success generating risk premiums that explain the empirical evidence.”[10]

4. Resolving the CSP

Economic data are generated by an unknown DGP. Economists develop models of these DGPs and test the validity of these models. Because economic models are approximations of the true DGP, tests based on such models potentially suffer from misspecification. Misspecification can cause distortions in both the size and power of tests of hypotheses and affect parameter estimates and other summary statistics depending on the nature and extent of the misspecification.

The resolution of the CSP presented here comes from noting that the estimates of β and λ need not be zero if the EH does not hold. Campbell and Shiller (1991, p. 500) allude to this possibility, but did not pursue it. The problem can be illustrated easily by assuming that long-term and short-term interest rates are generated by independent, stationary AR(1) processes. Given this assumption, it is easy to show that

$$E\hat{\beta} = \frac{(1/k)[(k-1)-\sum_{j=1}^{k-1}\theta^{j}]\delta}{1+\delta}, \qquad (10)$$

where θ is the autoregressive parameter for the short-term rate and δ is the ratio of the variance of the short-term rate to the long-term rate. Note that $\sum_{j=1}^{k-1}\theta^{j} < k-1$, so that (10) is strictly positive and increases in both k and δ.

On the other hand,

---

[9] Tzavalis and Wickens (1997) show that when their estimate of the risk premium is added to (6), the estimate of β is also insignificantly different from unity.

[10] Bekaert and Hodrick (2001, p. 1358).

$$E\hat{\lambda} = -((n-m)/m)[1/1+\delta] \tag{11}$$

is strictly negative. Moreover, the absolute value of $E\hat{\lambda}$ increases with k.[11]

These results stem from the fact that "the tail is wagging the dog": the results are driven by the fact that R_t^m appears with the same sign on both sides of (6) and R_t^n appears with opposite signs on both sides of (9).

4.1 *Generalizing the Bias of these Tests under the Alternative Hypothesis*

The actual results depend on the nature of the true DGP for long-term and short-term rates. Hence, it is useful to illustrate the problem more generally. This is done by noting that (4) can be rewritten as

$$(1/k)\sum_{i=0}^{k-1} R_{t+mi}^m - R_t^m = \alpha + \beta(R_t^n - R_t^m) + \beta' R_t^m + \omega_t, \tag{12}$$

where $\beta = 1 + \beta'$. Note that (12) reduces to (6) if and only if $\beta' = 0$. If (12) holds, but (6) is estimated, the expected value of the least-square estimator of β from (6) is

$$E\hat{\beta} = \beta + \beta' E\frac{\sum(\bar{R}_t^n - \bar{R}_t^m)\bar{R}_t^m}{\sum(\bar{R}_t^n - \bar{R}_t^m)^2} - E\frac{\sum \omega_t(\bar{R}_t^n - \bar{R}_t^m)}{\sum(\bar{R}_t^n - \bar{R}_t^m)^2}. \tag{13}$$

The second term on the right-hand side of (13) is zero if and only if the EH holds, i.e., $\beta' = 0$. When, $\beta' < 0$ (i.e., $\beta < 1$), the esimate of β will tend to be positive. It is easy to show that

$$\underset{N\to\infty}{P\lim}\,\hat{\beta} = \beta + \beta'\left[\frac{\sigma_{nm} - \sigma_m^2}{\sigma_n^2 - 2\sigma_{nm} + \sigma_m^2}\right]. \tag{14}$$

Noting that $\sigma_{nz} / \sigma_n^2 = \rho\delta^{1/2}$, this expression can be rewritten as

[11] The expression is slightly more complicated in the case where R_{t+m}^{n-m} does not exist and is replaced by R_{t+m}^n.

$$\underset{N\to\infty}{P\lim}\,\hat{\beta} = \beta + \beta'\left[\frac{\rho\delta^{1/2} - \delta}{1 - 2\rho\delta^{1/2} + \delta}\right], \tag{15}$$

where ρ is the coefficient of correlation between R_t^n and R_t^m. The estimate of β depends on both the extent to which β' differs from 0 and the term in brackets, which might be called the *relative variance factor*, *rvf*. In general, *rvf* can be either positive or negative; however, it is strictly negative whenever $\rho < \delta^{-1/2}$. Moreover, it is an increasing function of δ—other things the same, the estimate of β gets larger, the larger the variance of the short-term rate relative to the variance of the long-term rate. Hence, circumstances can easily arise where the estimate of β is positive when the EH does not hold.

Note that (4) can also be rewritten as

$$R_{t+m}^{n-m} - R_t^n = \mu + (m/(n-m))(R_t^n - R_t^m) + \beta'(n/n-m))R_t^n + \eta_t, \tag{16}$$

which reduces to (9) under the null hypothesis if and only if $\beta' = 0$. Moreover, if (9) is estimated but (16) holds,

$$\underset{N\to\infty}{P\lim}\,\hat{\lambda} = 1 + \beta'(n/m)\left[\frac{1 - \rho\delta^{1/2}}{1 - 2\rho\delta^{1/2} + \delta}\right]. \tag{17}$$

Note that the term in brackets is strictly positive under exactly the same circumstances where the bracketed term in (15) is strictly negative, i.e., whenever $\rho < \delta^{-1/2}$. Consequently, estimates of λ will tend to be less than 1 in exactly those circumstances where estimates of β will tend to be greater than zero. The difference is that the term in brackets in (17) is weighted by k, which can be large. Hence, estimates of λ can be negative.

The extent to which the conventional and contrarian tests generate results consistent with the CSP depends on two critical factors—$\rho < \delta^{-1/2}$ and β from (6) being less than one. The first condition is generally satisfied because short-term rates tend to be more variable than longer-term rates. Indeed, this is one of the implications of the EH.

The second condition can be shown to be satisfied whenever the EH does not hold exactly for any reason. To see why, assume that rather than being generated by (1), the long-term rate is generated by

$$R_t^n = (1/k)\sum_{i=0}^{k-1} E_t R_{t+mi}^m + \pi + \varphi Z_t. \tag{18}$$

This equation suggests that in addition to the market's expectation for the short-term rate, the long-term rate is determined by other factors, where Z_t denotes a vector of all other factors that determine the long-term rate, and $\varphi > 0$ denotes the response of the long-term rate to these factors. Using the same steps used to derive (6) from (1), (18) can be rewritten as

$$(1/k)\sum_{i=0}^{k-1} R_{t+mi}^m - R_t^m = \alpha + \beta^*(R_t^n - R_t^m) - \varphi Z_t + \omega_t. \tag{19}$$

Note that $\beta^* = 1$ is conditional on including Z_t. If (18) characterizes the DGP, but (6) rather than (19) is estimated,

$$\underset{N\to\infty}{P\lim}\,\hat{\beta} < 1. \tag{20}$$

This stems from the fact that by (18) R_t^n and Z_t are positively correlated. Note, however, that the same result would be obtained if R_t^n and Z_t were negatively related, i.e., if

$$R_t^n = (1/k)\sum_{i=0}^{k-1} E_t R_{t+mi}^m + \pi - \varphi Z_t. \tag{21}$$

The above results illustrate that if the long-term rate is determined by factors other than the market's expectation for the short-term rate, β implied by (6) will be strictly less than 1. Given the previous result in this subsection, this means that if the EH (equation 1) does not hold for any reason—a time-varying risk premium, the overreaction (or underreaction) of the long-term rate to changes in expectations for the short-term rate, market segmentation, etc.—tests of the EH using (6) and (9) are capable of generating results that are consistent with the CSP. The CSP can result from the facts that (a) the EH is not the true DGP processes for the long-term rate and (b) by design these tests can generate results that are consistent with the CSP when the EH is not the true DGP, whatever the reason.

It is important to note, however, that this does not necessarily imply that term spread is not useful for predicting the short-term rate or the short-term behavior of the long-term rate. Rather, because these tests are derived under a null hypothesis, if the EH is not the true DGP for the long-term rate, they need not reflect the predictability of the term spread. To do so, other factors that determine the long-term rate need to be identified and included in the specification of the test.

5. A Monte Carlo Investigation

In order to examine the extent to which these features of these tests can account for Campbell and Shiller's paradoxical results presented in Section 2, hypothetical "short-term" and "long-term" interest rates are generated where the EH does not hold, but many of the features of the historical data are preserved. The analysis in the previous section shows that the paradoxical results depend in part on the degree to which the EH is violated, with the CSP being strongest in cases where observed long-term rate is uncorrelated with the long-term rate implied by the EH. This section shows that, in most cases, Campbell and Shiller's results can be obtained using data where the EH cannot hold. This should not necessarily be taken as evidence that the long-term rates are unrelated to the market's expectation of the short-term rate.[12]

In any event, to see how these hypothetical long-term and short-term rates are obtained, consider the following VAR specification for the long-term and short-term interest rates

$$x_t = \sum_{h=1}^{L} \theta_h x_{t-h} + \varepsilon_t, \tag{22}$$

where $x_t = (R_t^n, R_t^m)'$, and ε_t is a 2 by 1 vector of *iid* independent shocks. The VAR specification is general enough to encompass a wide range of alternative DGPs. Moreover, it is easy to state restrictions on the VAR such that the EH does not hold.

To see how this is done, note that the EH implies the VAR satisfies the constraint

$$a_T(\Theta) \equiv e_1' - \frac{1}{k} e_2' \sum_{i=0}^{k-1} \sum_{h=1}^{L} \theta_{h,mi} = 0, \tag{23}$$

where e_1 and e_2 are elementary vectors that have 1 in the first and second position, respectively (e.g., Campbell and Shiller, 1987, 1991; Bekaert and Hodrick, 2001). The restrictions implied by (23) are highly nonlinear and can be tested using a LaGrange multiplier test (e.g., Bekaert and Hodrick, 2001). While it is difficult to determine a set of necessary conditions such that the EH will not hold, it is easy to show that $\theta_{h,21} = 0$ for all h is a sufficient condition, i.e., $\theta_{h,21} \neq 0$ is a necessary, but not a sufficient, condition for

[12] One possible explanation is that market participants have difficulty forecasting the future short-term rate beyond its current level. Diebold and Li (2003) recently document that many term structure models are unable to beat a random walk model at short horizons and only marginally improve on a random walk model at longer horizons.

the EH. This can be seen by noting that with the restriction $\theta_{h,21} = 0$ for all h, θ_h is an upper triangular matrix. The sum and products of upper triangular matrices are upper triangular. Hence, under this assumption, the restriction given by (23) requires $1 = 0$, i.e., the EH cannot hold.

The intuition for this result is straightforward. The EH indicates that today's long-term rate depends on today's expectation for tomorrow's short-term rate. If the EH holds, the short-term rate must be correlated with past values of the long-term rate. Any DGP where the short-term rate is independent of past values of the long-term rate is sufficient to guarantee that the EH does not hold. The reverse is not true, however. The existence of this correlation does not establish the EH. For example, past behavior of long-term rates might provide information about the future direction of interest rates—including short-term rates. Hence, short-term rates could be correlated with past long-term rates even if the EH does not hold.

5.1 ***The Monte Carlo Experiment***

To investigate the extent to which the proposed resolution of the CSP accounts for the results presented in Section 2, data are generated from (22) imposing the condition that $\theta_{h,21} = 0$. Specifically, (22) is estimated imposing the condition that $\theta_{h,21} = 0$ using the Campbell and Shiller (1991) data for the period 1952.01 – 1991.02.[13] The VAR was estimated for the 48 combinations of long-term and short-term rates where k is an integer. Following Bekaert and Hodrick (2001), the lag length was determined by the Schwartz

[13] There were six combinations of n and m for which the VAR was unstable when the restriction was imposed. These cases were (n, m) equal (2,1), (3,1), (4,1), (4,2), (48,24), and (120,60). In these instances it was also necessary to impose the restriction that the long-term rate was independent of lagged values of the short-term rate to obtain a stable model.

information criterion. In all but a few cases the optimal lag length was selected to be 2, so the lag length was taken to be 2.

Ten thousand samples of size 500 and 1000 are generated from the model using an *iid* bootstrap of the residuals obtained from the estimated VAR. The bootstrap procedure preserves the contemporaneous correlation, skewness, and kurtosis in the residuals. In each experiment, the initial 1,000 observations were discarded to minimize the effect of the initial conditions. These data were used to estimate (6) and (9). In the instances where R_{t+im}^{n-m} is not available, the usual procedure of replacing R_{t+im}^{n-m} with R_{t+im}^{n} is followed. In cases where R_{t+im}^{n-m} is available and R_{t+im}^{n-m}, R_t^n and R_t^m are different, the data are generated using a three variable VAR. The estimated standard errors were obtained using the Newey-West (1987) procedure.

5.2. *The Monte Carlo Results*

As noted earlier, a necessary but not sufficient condition for the EH is that coefficients on the lagged long-term rate in the short-term rate equation not equal zero. In every case the restriction that the coefficients on lagged values of the long-term rate in the short-term rate equation are zero was rejected at a low significance level. Hence, the data satisfy the necessary conditions for the EH.[14]

Despite the fact that the hypothesis $\theta_{1,21} = \theta_{2,21} = 0$ is easily rejected for all specifications, data generated from the VAR with these restrictions imposed do a reasonably good job of characterizing the historical data. Table 3 reports the mean, median, and standard deviations of the actual data and the hypothetical data generated by

[14] Dittmar and Thornton (2003) test the restriction given by (23) using these data and the LaGrange multiplier test of Bekaert and Hodrick (2001). These restrictions are rejected at the 5 percent significance level or lower for maturity pairing except at the long end of the maturity spectrum.

the VAR models. The VARs were estimated for all combinations of n and m, such that k is integer valued; however, the results were relatively insensitive to choice of n and m. Hence, to conserve space, only the statistics for $m=1$ are presented in Table 3.

As was the case for Bekaert et al. (2001), the means are somewhat larger for the model data and the kurtosis is somewhat less than in the U.S. data. The standard deviations are also somewhat smaller, but like the actual data, they tend to be very similar for maturities up to about 36 months and then decline. All in all, the model generates data that appear to approximate the historical data reasonably well.

Tables 4 and 5 report the results estimating the conventional test using artificial data for sample sizes 500 and 1000, respectively. Tables 6 and 7 report the analogous results for the contrarian test. The first entry in each cell is the average estimate of the parameter (β or λ), the second is the range of the parameter estimates, the third is the average estimate of $\overline{R}^2$, the fourth is the percentage of times that the null hypothesis that the estimated coefficient is zero is rejected, and the fifth is the percentage of times that the EH is rejected, i.e., the power of the test.

The results in Tables 4 and 5 suggest that the estimates of β are nearly always positive. The estimates are quite small for small values of k, but increase as k gets larger. The range of parameter estimates tends to be large and becomes larger, the larger is k. Both the average estimate of β and the range decrease as the sample size increases; however, estimates of β appear to decline slowly so that the frequency of rejecting the hypothesis $\beta=0$ increases. The power of the test is very high when the short-term rate is less than three years, but generally declines as k increases and is low when n=120. As expected, the power of the test improves as the sample size increases.

The results presented in Tables 6 and 7 show that the estimates of λ are nearly always negative. Moreover, consistent with the analysis in Section 4, the estimates tend to be more negative as k increases. As was the case for the conventional test, the absolute value of the estimates of λ decline as the sample size increases, albeit slowly. Hence, even for samples of 1000, negative estimates of λ less than -1 are common. The power of the test is very high for all interest rate pairs, save those with long maturities, and increases with the sample size.

These results show that the conventional and contrarian tests tend to generate results that are consistent with the CSP when the EH does not hold: estimates of β are nearly always positive, while estimates of λ are nearly always negative.

To assess the extent to which the explanation advanced in Section 4 accounts for the CSP, Tables 8 and 9 report the probably that the parameter estimates obtained with actual data could have been generated by a DGP where the EH does not hold. Specifically, the probabilities $1-\Pr(\tilde{\varphi} \leq |\overline{\varphi}-\hat{\varphi}|)$ are reported, where $\tilde{\varphi}$ is an estimate of φ (β or λ) obtained from the hypothetical data, $\overline{\varphi}$ is the mean estimate of these parameters obtained from the Monte Carlo experiments, and $\hat{\varphi}$ is the estimate of these parameters reported in Tables 1 and 2.

These results are reported only for sample size 500 because this sample size more nearly matches the sample sizes used to generate the results in Tables 1 and 2.[15] All of the instances where $1-\Pr(\tilde{\varphi} \leq |\overline{\varphi}-\hat{\varphi}|)$ is greater than or equal to 5 percent are highlighted. This resolution of the CSP does a better job accounting for the smirk than the smile. This is illustrated in Figures 3 and 4, which present the smile (for $m=1$) obtained from Tables 1

[15] The largest sample size is 469 observations.

and 4, respectively, and the smirk obtained from Tables 2 and 6, respectively. It is clear from Figure 3 that this resolution of the CSP does not account for estimates of β when the short-term rate has a maturity of 9 months or less. It also does not account for estimates of β when the long-term rate has a maturity of 120 months. In both cases, estimates of β from the hypothetical data are too small relative to the estimates obtained using historical data.

Despite these limitations, using a 5 percent significance level as the criterion, the resolution of the CSP presented here accounts for the results for 34 of the 48 estimates of β reported in Table 1 and for 40 of the 48 estimates of λ in Table 2. Hence, with relatively few exceptions, the Monte Carlo experiments show that the explanation advanced in Section 2 resolves the CSP.

5. Conclusions and Implications

The expectations hypothesis of the term structure asserts that long-term rates are determined by the market's expectation of the short-term rate over the holding period of the long-term rate plus a constant risk premium. One of the most puzzling results in this literature is the CSP—the fact that regressions of the long-term change in the short-term rate on the spread between the long-term and short-term rates generate results that are favorable to the EH in that the term spread predicts the correct direction of the future change in the short-term rate, while an equivalent test under the null hypothesis that regresses short-term changes in the long-term rate on the term spread generates results that are unfavorable to the EH in that the term spread predicts the wrong direction of the change in the long-term rate.

Previous attempts to resolve the CSP have met with limited success. Researchers have only been able to account for the CSP using statistical models that are flexible enough to capture the observed variation in the term premiums in the data. The resolution offered here is based on the observation that these tests tend to generate results consistent with the CSP when the EH does not represent the true data generating process for the long-term rate. This tendency arises from the fact that, for one test, the current-period short-term rate appears symmetrically on both sides of the equation, while for the other test, the current long-term rate appears on both sides of the equation with opposite signs. Consequently, for the one test the slope coefficient will be biased toward 1 when the EH does not hold, while for the other test the estimated slope coefficient may be negative.

Monte Carlo experiments show that this explanation goes a long way toward resolving the CSP. For most combinations of the long-term and short-term rates, the results from these tests using historical data could have been generated by applying these tests to data for which the EH does not hold.

There are several implications of these results. The first is that there is really no paradox. By their design these tests tend to generate results consistent with the CSP when the EH does not hold. Given the widespread failure of the EH, the proposition that the EH does not hold can hardly be considered controversial. While analysts are understandably interested in knowing exactly why the EH does not hold, the fact that it does not hold is sufficient to account for most of Campbell and Shiller's (1991) paradoxical results.

A second implication is that the magnitude of the estimated slope coefficients and the adjusted R-squares from these tests are not useful for assessing the validity of the EH when the EH is rejected by these tests. Given that the estimates from the contrarian test were frequently negative, this has long been assumed to be true of it. The analysis presented here shows that this conclusion is equally true for the conventional test. This is important because the size of the estimated slope coefficient from the contrarian test and the size of the adjusted R-square have been used an indicator of the potential validity of the EH, even when the EH is rejected. This implication is made more important by the fact that similar tests are used to test other important financial theories.

A third implication comes from combining the results of Bekaert et al. (2001) with the results presented here. Bekaert et al. demonstrate that the conventional and contrarian tests of the EH are severely positively biased in small samples when the EH holds. This bias is due to the extreme persistence in the short-term rate under the null hypothesis and remains even in a relatively large sample. As a result, they find that the evidence against the EH is much stronger using their small-sample distributions than the asymptotic distributions. The analysis presented here complements this work by showing that these tests are biased in the direction of the CSP when the EH does not hold. As in Bekaert et al., (1997), these biases remain relatively severe even in very large samples.

Fourth, because the estimated coefficients of these equations can be severely biased away from their theoretical values under both the null and the alternative hypotheses, researchers might do well to consider alternative procedures for testing the EH. One such procedure would be to test the EH by estimating a general VAR for interest rates and imposing the restrictions imposed by the EH. This procedure was proposed by Campbell and Shiller (1987) and has recently been made operational by Bekaert and Hodrick (2001). This test is less sensitive to the bias noted by Bekaert et al. (1997), and, because VARs are able to capture comovements among rates that cannot be captured in single equations, this test should be more powerful than the conventional and contrarian tests derived under the null hypothesis. Moreover, this methodology is easily extended to models of more than two interest rates or that include other economic variables, such as output and inflation, and can be applied under the assumption that interest rates are stationary or integrated of order one, but cointegrated (e.g., Campbell and Shiller, 1987; Dittmar and Thornton, 2003).

Finally, the results here show that the CSP does not necessarily have implications for the predictive power of the term spread for either long-term changes in the short-term rate or short-term changes in the long-term rate. Rather, they suggest that if the EH does not represent the true DGP for the long-term rate, the predictive power of the term spread can be identified only if the other factors that determine the long-term rate are identified and included in the specification of the test.

REFERENCES:

Backus, David K., Allen W. Gregory, and Stanley E. Zin. (1989). "Risk Premiums in the Term Structure: Evidence from Artificial Economies." *Journal of Monetary Economics* 24, 371-99.

Baillie, Richard T., and Tim Bollerslev. (2000). "The Forward Premium Anomaly is Not as Bad as You Think." *Journal of International Money and Finance* 19, 471-88.

Balduzzi, Pierluigi, Guiseppe Bertola, and Silverio Foresi. (1997). "A Model of Target Changes and the Term Structure of Interest Rates." *Journal of Monetary Economics* 39, 223-49.

Bansal, Ravi, and Hao Zhou. (2002). "Term Structure of Interest Rates with Regime Shifts." *Journal of Finance* 57, 1997-2043.

Bekaert, Geert, Robert J. Hodrick, and David A. Marshall. (2001). "Peso Problem Explanations for Term Structure Anomalies." *Journal of Monetary Economics* 48, 241-70.

Bekaert, Geert, Robert J. Hodrick, and David A. Marshall. (1997). "On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates." *Journal of Financial Economics* 44, 309-48.

Bekaert, Geert, and Robert J. Hodrick. (2001). "Expectations Hypotheses Tests." *Journal of Finance* 56, 1357-94.

Campbell, John Y. (1995). "Some Lessons from the Yield Curve." *Journal of Economic Perspectives* 9, 129-52.

Campbell, John Y., and Robert J. Shiller. (1987). "Cointegration and Tests of Present Value Models." *Journal of Political Economy* 95, 1062-88.

Campbell, John Y., and Robert J. Shiller. (1991). "Yield Spreads and Interest Rate Movements: A Bird's Eye View." *Review of Economic Studies* 58, 495-514.

Chen, Ren-Raw, and Louis O. Scott. (1992). "Pricing Interest Rate Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure." *Review of Financial Studies* 5, 613-36.

Chen, Ren-Raw, and Louis O. Scott. (1993). "Maximum Likelihood Estimation for a Multi-Factor Equilibrium Model of the Term Structure of Interest Rates." *Journal of Fixed Income* 4, 14-31.

Cox, John C., Johnathan E. Ingersoll, Jr., and Stephen A. Ross. (1985). "A Theory of the Term Structure of Interest Rates." *Econometrica* 53, 385-407.

Dai, Qiang, and Kenneth J. Singleton. (2002). "Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure." *Journal of Financial Economics* 63, 415-441.

Diebold, Francis X., and Canlin Li. (2003). "Forecasting the Term Structure of Government Bond Yields." National Bureau of Economic Research Working Paper No. 10048.

Dittmar, Robert D., and Daniel L. Thornton. (2003). "New Evidence on the Expectations Hypothesis of the Term Structure of Bond Yields." Federal Reserve Bank of St. Louis Working Paper 2003-021B.

Dotsey, Michael, and Christopher Otrok. (1995). "The Rational Expectations Hypothesis of the Term Structure, Monetary Policy, and Time-Varying Term Premia." Federal Reserve Bank of Richmond *Economic Quarterly* 81, 65-81.

Driffill, John, Zacharias Psaradakis, and Martin Sola. (1997). "A Reconciliation of Paradoxical Empirical Results on the Expectations Model of the Term Structure." *Oxford Bulletin of Economics and Statistics* 59, 29-42.

Duffee, Gregory R., (2002). "Term Premia and Interest Rate Forecasts In Affine Models." *Journal of Finance* 57, 405-443.

Fama, Eugene F. (1984). "The Information in the Term Structure." *Journal of Financial Economics* 13, 509-28.

Froot, Kenneth A. (1989). "New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates." *Journal of Finance* 44, 283-305.

Goodfriend, Marvin. (1991). "Interest Rates and the Conduct of Monetary Policy." *Carnegie-Rochester Conference Series on Public Policy* 34, 7-30.

Hansen, Lars P., and Robert J. Hodrick. (1980). "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis." *Journal of Political Economy* 88, 829-53.

Hardouvelis, Gikas A. (1994). "The Term Structure Spread and Future Changes in Long and Short Rates in the G7 Countries: Is There a Puzzle?" *Journal of Monetary Economics* 33, 255-83.

Hardouvelis, Gikas A. (1988). "The Predictive Power of the Term Structure During Recent Monetary Regimes." *Journal of Finance* 43, 339-56.

Hsu, Chiente, and Peter Kugler. (1997). "The Revival of the Expectations Hypothesis of the US Term Structure of Interest Rates." *Economic Letters* 55, 115-20.

Kozicki, Sharon, and P. A. Tinsely. (2001). “What Do You Expect? Imperfect Policy Credibility and Tests of the Expectations Hypothesis.” Federal Reserve Bank of Kansas City working paper, RWP 01/02. Forthcoming in the *Journal of Monetary Economics*.

Longstaff, Francis A. (2000). “The Term Structure of Very Short-Term Rates: New Evidence for the Expectations Hypothesis.” *Journal of Financial Economics* 58, 397-415.

Mankiw, Gregory N., and Jeffrey A. Miron. (1986). “The Changing Behavior of the Term Structure of Interest Rates.” *Quarterly Journal of Economics* 101, 211-28.

McCallum, Bennett T. (1994a). “Monetary Policy and the Term Structure of Interest Rates.” National Bureau of Economic Research Working Paper No. 4938.

McCallum, Bennett T. (1994b). “A Reconsideration of the Uncovered Interest Parity Relationship.” *Journal of Monetary Economics* 33, 105-32.

McCulloch, J. Huston. (1990) “U.S. Government Term Structure Data, 1947 – 1987.” In *The Handbook of Monetary Economics*, edited by Benjamin M. Friedman and Frank H. Hahn, pp 672 - 715. Amsterdam: North Holland.

McCulloch, J. Huston, and Heon-Chul Kwon. (1993). “U.S. Term Structure Data, 1947-1991.” Ohio State University Working Paper #93-6.

Newey, Whitney K., and Kenneth D. West. (1987). “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica* 55, 703-8.

Roberds, William, David Runkle, and Charles H. Whiteman. (1996). “A Daily View of Yield Spreads and Short-Term Interest Rate Movements.” *Journal of Money, Credit and Banking* 28, 34-53.

Roberds, William, and Charles H. Whiteman. (1999). “Endogenous Term Premia and Anomalies in the Term Structure of Interest Rates: Explaining the Predictability Smile.” *Journal of Monetary Economics* 44, 555-80.

Rudebusch, Glenn D. (1995). “Federal Reserve Interest Rate Targeting, Rational Expectations, and the Term Structure.” *Journal of Monetary Economics* 35, 245-74.

Shiller, Robert J., John Y. Campbell, and Kermit L. Schoenholtz. (1983). “Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates.” *Brookings Papers on Economic Activity* (1), 173-217.

Simon, David P. (1990). “Expectations and the Treasury Bill-Federal Funds Rate Spread over Recent Monetary Policy Regimes.” *Journal of Finance* 45, 467-77.

Stambaugh, Robert F. (1988). “The Information in Forward Rates: Implications for Models of the Term Structure.” *Journal of Financial Economics* 21, 41-70.

Tzavalis, Elias, and Michael R. Wickens. (1997). “Explaining the Failure of the Term Spread Models of the Rational Expectations Hypothesis of the Term Structure.” *Journal of Money, Credit and Banking* 29, 364-80.

Woodford, Michael. (2001). “Monetary Policy in the Information Economy.” In *Economic Policy for the Information Economy*. Kansas City, Federal Reserve Bank of Kansas City.

Table 1: Estimates from the Conventional Test $(1/k)\sum_{i=0}^{k-1} r_{t+im}^{m} - r_{t}^{m} = \alpha + \beta(r_{t}^{n} - r_{t}^{m}) + v_{t}$

n/m	1	2	3	4	6	9	12	24	60
2	0.502								
	0.107								
	5.247								
	5.213								
3	0.467								
	0.083								
	3.166								
	3.607								
4	0.471	0.224							
	0.073	0.008							
	2.484	0.926							
	2.795	3.217							
6	0.320	0.087	-0.075						
	0.031	0.000	-0.001						
	2.190	0.574	0.414						
	4.656	5.988	5.906						
9	0.254		0.061						
	0.019		-0.001						
	1.649		0.432						
	4.842		6.602						
12	0.272	0.157	0.114	0.023	0.093				
	0.022	0.005	0.001	-0.002	-0.001				
	1.308	0.852	0.629	0.128	0.308				
	3.497	4.583	4.897	5.423	2.992				
24	0.363	0.337	0.318	0.280	0.238		0.089		
	0.038	0.029	0.024	0.016	0.010		-0.001		
	1.632	1.523	1.321	1.038	0.748		0.252		
	2.858	2.992	2.839	2.667	2.398		2.590		
36	0.401	0.428	0.426	0.404	0.369	0.283	0.209		
	0.048	0.050	0.046	0.039	0.030	0.016	0.006		
	1.422	1.500	1.465	1.318	1.123	0.774	0.492		
	2.129	2.006	1.974	1.945	1.924	1.959	1.860		
48	0.443	0.505	0.522	0.515	0.505		0.375	0.090	
	0.055	0.068	0.069	0.064	0.059		0.025	-0.001	
	1.151	1.339	1.436	1.436	1.482		1.150	0.150	
	1.450	1.312	1.314	1.353	1.452		1.919	1.523	
60	0.511	0.583	0.607	0.607	0.612		0.558		
	0.076	0.095	0.097	0.092	0.089		0.059		
	1.230	1.387	1.461	1.473	1.594		1.845		
	1.177	0.990	0.946	0.952	1.010		1.462		
120	1.402	1.473	1.507	1.523	1.567		1.618	1.759	2.290
	0.690	0.721	0.729	0.711	0.687		0.550	0.399	0.196
	9.552	10.595	10.796	10.316	9.393		8.318	5.899	2.824
	2.738	3.402	3.633	3.541	3.399		3.178	2.545	1.591

Table 2: Estimates from the Contrarian Test $r_{t+mi}^{n-m} - r_t^n = \mu + \lambda[m/(n-m)(r_t^n - r_t^m) + \varpi_t$

n/m	1	2	3	4	6	9	12	24	60
2	0.003								
	-0.002								
	0.017								
	5.213								
3	-0.145								
	-0.001								
	0.407								
	3.221								
4	-0.346	-0.553							
	0.003	0.013							
	0.644	1.146							
	2.507	3.217							
6	-0.493	-1.124	-1.151						
	0.004	0.038	0.056						
	0.954	2.586	3.160						
	2.890	4.888	5.906						
9	-0.915		-1.465						
	0.009		0.053						
	1.747		3.044						
	3.656		5.121						
12	-1.029	-1.325	-1.728	-1.130	-0.814				
	0.008	0.024	0.055	0.031	0.020				
	1.763	3.306	3.388	2.689	1.342				
	3.475	5.802	5.348	5.069	2.992				
24	-1.448	-1.665	-1.471	-1.259	-0.718		-0.822		
	0.008	0.021	0.024	0.023	0.011		0.024		
	1.875	2.476	2.615	2.676	2.076		1.169		
	3.170	3.963	4.393	4.801	4.965		2.590		
36	-1.890	-1.937	-1.708	-1.465	-0.983	-0.961	-1.134		
	0.011	0.022	0.025	0.024	0.016	0.026	0.035		
	10.053	4.945	6.235	1.713	1.932	1.934	1.492		
	15.372	7.498	9.885	2.882	3.898	3.946	2.808		
48	-2.264	-2.306	-1.973	-1.706	-1.210		-1.426	-0.821	
	0.012	0.024	0.026	0.026	0.020		0.044	0.014	
	2.623	2.860	2.597	2.443	1.884		2.560	0.686	
	3.781	4.100	3.913	3.875	3.440		4.356	1.523	
60	-2.613	-2.621	-2.237	-1.969	-1.462		-1.680		
	0.013	0.026	0.028	0.029	0.024		0.050		
	2.919	2.801	2.905	3.497	4.004		1.637		
	4.036	3.869	4.204	5.273	6.743		2.612		
120	-4.220	-4.068	-3.693	-3.316	-2.629		-2.451	-1.241	3.581
	0.018	0.032	0.039	0.042	0.038		0.051	0.014	0.129
	5.504	4.978	3.958	3.759	3.362		4.877	1.553	2.207
	6.808	6.201	5.030	4.892	4.641		6.867	2.804	1.591

Table 3: Summary Statistics for U.S. and Model Data

	Mean		Median		S.D.	
Term	U.S.	Model	U.S.	Model	U.S.	Model
1	5.31	5.69	4.94	5.73	3.06	2.67
2	5.51	5.91	5.15	6.00	3.12	2.71
3	5.64	6.06	5.27	6.14	3.14	2.73
4	5.74	6.17	5.34	6.22	3.15	2.71
6	5.88	6.27	5.55	6.32	3.18	2.71
9	6.00	6.39	5.74	6.44	3.18	2.72
12	6.08	6.48	5.82	6.52	3.17	2.69
24	6.27	6.72	6.15	6.76	3.12	2.58
36	6.39	6.85	6.37	6.87	3.08	2.47
48	6.47	6.98	6.50	7.00	3.07	2.41
60	6.53	7.07	6.53	7.09	3.06	2.36
120	6.68	7.30	6.56	7.32	3.01	2.21

Table 4: Monte Carlo Estimates of the Conventional Test—Sample Size 500

	1	2	3	4	6	9	12	24	60
2	0.006								
	-0.059, 0.065								
	0.000								
	0.060								
	1.000								
3	0.014								
	-0.062, 0.118								
	0.002								
	0.098								
	1.000								
4	0.022	0.012							
	-0.066, 0.142	-0.114, 0.167							
	0.005	0.001							
	0.120	0.048							
	1.000	1.000							
6	0.024	-0.020	-0.046						
	-0.257, 0.317	-0.449, 0.464	-0.589, 0.424						
	0.005	0.004	0.003						
	0.078	0.075	0.088						
	1.000	1.000	1.000						
9	0.033		-0.019						
	-0.352, 0.484		-0.708, 0.641						
	0.008		0.005						
	0.084		0.075						
	1.000		1.000						
12	0.063	0.036	0.029	0.027	0.038				
	-0.374, 0.624	-0.510, 0.644	-0.623, 0.686	-0.623, 0.712	-0.533, 0.716				
	0.012	0.009	0.008	0.007	0.007				
	0.114	0.083	0.084	0.077	0.071				
	1.000	1.000	1.000	0.998	0.999				
24	0.172	0.160	0.155	0.157	0.168		0.109		
	-0.533, 0.876	-0.656, 0.838	-0.660, 1.029	-0.658, 1.112	-0.626, 0.963		-0.457, 0.821		
	0.036	0.027	0.024	0.024	0.027		0.022		
	0.285	0.218	0.198	0.195	0.210		0.144		
	0.998	0.991	0.984	0.977	0.981		0.997		
36	0.278	0.270	0.271	0.279	0.292	0.275	0.263		
	-0.500, 1.040	-0.594, 1.198	-0.696, 1.353	-0.839, 1.212	-0.729, 1.249	-0.555, 1.231	-0.633, 1.314		
	0.064	0.053	0.049	0.048	0.055	0.059	0.054		
	0.455	0.379	0.357	0.347	0.377	0.394	0.357		
	0.971	0.935	0.916	0.888	0.881	0.922	0.915		
48	0.352	0.347	0.347	0.361	0.381		0.352	0.263	
	-0.456, 1.298	-0.615, 1.286	-0.738, 1.360	-0.737, 1.534	-0.588, 1.456		-0.631, 1.489	-0.607, 1.681	
	0.089	0.076	0.071	0.072	0.083		0.083	0.052	
	0.544	0.484	0.458	0.465	0.501		0.483	0.326	
	0.918	0.867	0.837	0.794	0.784		0.827	0.856	
60	0.419	0.417	0.418	0.431	0.448		0.432		
	-0.526, 1.394	-0.679, 1.331	-0.718, 1.411	-0.770, 1.497	-1.022, 1.715		-0.833, 1.620		
	0.117	0.101	0.095	0.096	0.107		0.110		
	0.633	0.571	0.550	0.545	0.580		0.581		
	0.831	0.773	0.742	0.706	0.707		0.736		
120	0.625	0.639	0.644	0.657	0.688		0.681	0.643	0.389
	-0.586, 1.628	-1.104, 1.718	-1.008, 1.724	-1.033, 1.859	-0.810, 1.708		-0.821, 1.856	-1.016, 2.205	-1.218, 2.370
	0.225	0.208	0.201	0.203	0.227		0.233	0.214	0.117
	0.816	0.784	0.761	0.764	0.797		0.788	0.739	0.526
	0.518	0.456	0.430	0.410	0.385		0.447	0.516	0.707

Table 5: Monte Carlo Estimates of the Conventional Test—Sample Size 1000

	1	2	3	4	6	9	12	24	60
2	0.005								
	-0.027, 0.042								
	0.000								
	0.063								
	1.000								
3	0.012								
	-0.037, 0.065								
	0.002								
	0.137								
	1.000								
4	0.018	0.010							
	-0.046, 0.087	-0.068, 0.106							
	0.003	0.001							
	0.164	0.041							
	1.000	1.000							
6	0.022	-0.021	-0.044						
	-0.158, 0.217	-0.305, 0.278	-0.415, 0.320						
	0.003	0.002	0.002						
	0.080	0.065	0.091						
	1.000	1.000	1.000						
9	0.030		-0.019						
	-0.268, 0.321		-0.443, 0.587						
	0.004		0.003						
	0.076		0.066						
	1.000		1.000						
12	0.059	0.033	0.027	0.022	0.036				
	-0.255, 0.375	-0.398, 0.473	-0.484, 0.557	-0.475, 0.506	-0.390, 0.529				
	0.007	0.005	0.004	0.003	0.004				
	0.119	0.075	0.071	0.069	0.076				
	1.000	1.000	1.000	1.000	1.000				
24	0.164	0.152	0.148	0.149	0.156		0.092		
	-0.267, 0.610	-0.453, 0.686	-0.344, 0.744	-0.412, 0.779	-0.360, 0.677		-0.248, 0.552		
	0.026	0.018	0.016	0.015	0.018		0.014		
	0.370	0.264	0.236	0.229	0.251		0.152		
	1.000	1.000	1.000	1.000	1.000		1.000		
36	0.260	0.256	0.256	0.262	0.272	0.248	0.221		
	-0.291, 0.792	-0.472, 0.867	-0.396, 0.948	-0.445, 0.920	-0.383, 0.942	-0.295, 0.986	-0.233, 1.019		
	0.049	0.039	0.036	0.035	0.041	0.043	0.037		
	0.578	0.476	0.436	0.430	0.469	0.477	0.410		
	1.000	0.998	0.994	0.991	0.991	0.996	0.996		
48	0.333	0.330	0.331	0.339	0.353		0.306	0.204	
	-0.167, 0.949	-0.275, 0.975	-0.377, 1.010	-0.413, 1.007	-0.381, 1.096		-0.282, 1.040	-0.434, 1.090	
	0.071	0.059	0.055	0.054	0.063		0.059	0.031	
	0.704	0.615	0.580	0.578	0.616		0.573	0.326	
	0.996	0.986	0.979	0.969	0.957		0.976	0.986	
60	0.395	0.395	0.399	0.404	0.421		0.383		
	-0.366, 0.974	-0.527, 1.108	-0.352, 1.113	-0.385, 1.114	-0.327, 1.132		-0.281, 1.260		
	0.093	0.080	0.075	0.074	0.085		0.082		
	0.781	0.712	0.690	0.674	0.724		0.685		
	0.983	0.959	0.941	0.919	0.917		0.930		
120	0.588	0.602	0.606	0.613	0.640		0.615	0.536	0.279
	-0.259, 1.248	-0.297, 1.392	-0.260, 1.276	-0.370, 1.355	-0.431, 1.324		-0.303, 1.560	-0.399, 1.585	-0.646, 1.569
	0.182	0.167	0.159	0.158	0.181		0.179	0.150	0.064
	0.933	0.911	0.891	0.883	0.910		0.890	0.803	0.491
	0.791	0.698	0.656	0.628	0.596		0.644	0.733	0.917

Table 6: Monte Carlo Estimates of the Contrarian Test—Sample Size 500

	1	2	3	4	6	9	12	24	60
2	-0.019								
	-0.153, 0.073								
	0.001								
	0.094								
	1.000								
3	-0.641								
	-1.070, -0.206								
	0.373								
	1.000								
	1.000								
4	-0.390	-0.032							
	-1.080, 0.285	-0.315, 0.203							
	0.122	0.002							
	0.910	0.067							
	1.000	1.000							
6	-0.522	-0.209	-0.804						
	-1.186, 0.023	-1.755, 2.048	-1.934, 0.130						
	0.029	0.009	0.044						
	0.941	0.137	0.857						
	1.000	0.786	1.000						
9	-1.050		-1.726						
	-2.171, -0.149		-3.343, -0.213						
	0.043		0.106						
	0.987		0.991						
	1.000		1.000						
12	-1.178	-1.385	-2.315	-1.381	-0.766				
	-2.416, -0.066	-2.762, -0.021	-4.010, -0.690	-2.855, 0.102	-1.901, 0.534				
	0.041	0.060	0.158	0.076	0.058				
	0.985	0.966	0.999	0.928	0.720				
	1.000	1.000	1.000	0.999	0.999				
24	-1.566	-1.742	-1.811	-1.657	-0.983		-0.134		
	-3.344, 0.043	-3.991, 0.021	-4.029, 0.360	-4.227, 0.501	-2.624, 1.067		-1.255, 1.105		
	0.034	0.054	0.072	0.074	0.058		0.020		
	0.966	0.954	0.942	0.896	0.648		0.128		
	1.000	0.999	0.999	0.997	0.988		0.980		
36	-1.936	-2.014	-1.982	-1.751	-1.025	-0.459	-0.422		
	-4.375, -0.097	-4.417, 0.034	-4.411, 0.412	-4.103, 1.069	-3.485, 1.365	-2.421, 1.443	-2.684, 2.020		
	0.033	0.051	0.064	0.063	0.045	0.029	0.038		
	0.967	0.939	0.911	0.829	0.508	0.236	0.265		
	1.000	0.998	0.995	0.987	0.949	0.902	0.850		
48	-2.310	-2.383	-2.310	-2.033	-1.260		-0.521	0.040	
	-4.870, 0.551	-4.971, 0.407	-5.163, 0.280	-4.848, 0.933	-3.772, 1.536		-3.195, 2.365	-1.841, 2.237	
	0.034	0.055	0.068	0.067	0.051		0.037	0.032	
	0.969	0.952	0.925	0.847	0.564		0.263	0.191	
	0.999	0.997	0.995	0.986	0.938		0.821	0.757	
60	-2.652	-2.700	-2.587	-2.305	-1.524		-0.609		
	-5.436, -0.261	-5.393, -0.020	-6.081, 0.464	-5.351, 1.751	-4.662, 1.515		-3.063, 2.298		
	0.035	0.058	0.071	0.072	0.059		0.037		
	0.975	0.964	0.930	0.865	0.635		0.277		
	0.999	0.997	0.994	0.985	0.940		0.776		
120	-4.255	-4.163	-3.952	-3.540	-2.542		-1.307	-0.339	0.014
	-8.940, 0.153	-8.559, -0.207	-8.227, 0.592	-7.621, 0.886	-6.010, 2.404		-5.218, 3.622	-3.750, 3.898	-3.117, 3.414
	0.042	0.067	0.083	0.086	0.078		0.057	0.042	0.081
	0.991	0.977	0.959	0.910	0.735		0.435	0.262	0.408
	0.999	0.997	0.994	0.980	0.928		0.776	0.603	0.690

Table 7: Monte Carlo Estimates of the Contrarian Test—Sample Size 1000

	1	2	3	4	6	9	12	24	60
2	-0.016								
	-0.085, 0.037								
	0.001								
	0.121								
	1.000								
3	-0.641								
	-0.990, -0.332								
	0.393								
	1.000								
	1.000								
4	-0.387	-0.029							
	-0.907, 0.059	-0.193, 0.131							
	0.122	0.001							
	0.983	0.074							
	1.000	1.000							
6	-0.512	-0.136	-0.795						
	-0.928, -0.177	-1.311, 1.660	-1.516, -0.088						
	0.028	0.005	0.042						
	0.998	0.121	0.982						
	1.000	0.899	1.000						
9	-1.036		-1.720						
	-1.811, -0.456		-2.891, -0.456						
	0.043		0.105						
	1.000		1.000						
	1.000		1.000						
12	-1.160	-1.372	-2.308	-1.379	-0.762				
	-1.928, -0.454	-2.362, -0.510	-3.637, -1.147	-2.408, -0.368	-1.528, 0.102				
	0.040	0.059	0.157	0.075	0.055				
	1.000	1.000	1.000	0.995	0.929				
	1.000	1.000	1.000	1.000	1.000				
24	-1.534	-1.714	-1.777	-1.638	-0.972		-0.128		
	-2.659, -0.407	-3.290, -0.386	-3.282, -0.420	-3.040, -0.299	-2.148, 0.248		-0.876, 0.609		
	0.033	0.053	0.069	0.072	0.054		0.012		
	1.000	0.999	0.998	0.992	0.894		0.128		
	1.000	1.000	1.000	1.000	1.000		1.000		
36	-1.898	-1.982	-1.952	-1.730	-1.014	-0.456	-0.454		
	-3.447, -0.531	-3.421, -0.446	-3.598, -0.276	-3.439, 0.091	-2.614, 0.393	-1.655, 0.900	-1.814, 1.161		
	0.032	0.050	0.062	0.060	0.041	0.023	0.031		
	0.999	0.998	0.994	0.979	0.779	0.350	0.376		
	1.000	1.000	1.000	1.000	0.999	0.995	0.984		
48	-2.263	-2.346	-2.273	-2.007	-1.252		-0.533	-0.007	
	-4.013, -0.748	-4.304, -0.539	-3.907, -0.371	-3.953, -0.108	-3.171, 0.340		-1.926, 1.002	-1.199, 1.358	
	0.033	0.054	0.066	0.065	0.048		0.029	0.014	
	0.999	0.999	0.996	0.984	0.825		0.364	0.125	
	1.000	1.000	1.000	1.000	0.999		0.972	0.946	
60	-2.600	-2.649	-2.561	-2.274	-1.499		-0.618		
	-4.728, -0.971	-4.469, -0.912	-4.599, -0.622	-4.540, 0.377	-3.878, 0.351		-2.192, 1.241		
	0.035	0.056	0.069	0.070	0.055		0.029		
	1.000	0.999	0.998	0.986	0.887		0.370		
	1.000	1.000	1.000	1.000	0.999		0.952		
120	-4.180	-4.102	-3.894	-3.498	-2.535		-1.324	-0.406	-0.061
	-6.890, -1.583	-6.911, -0.495	-6.790, -0.917	-6.311, -0.747	-5.237, 0.399		-3.712, 1.116	-2.309, 2.250	-1.812, 2.328
	0.041	0.066	0.081	0.084	0.075		0.050	0.026	0.037
	1.000	1.000	0.999	0.995	0.945		0.634	0.261	0.315
	1.000	1.000	1.000	1.000	0.997		0.956	0.813	0.881

Table 8: Probability Estimates of β Came From Hypothetical Data									
n\m	1	2	3	4	6	9	12	24	60
2	.000								
3	.000								
4	.000	.000							
6	.008	.314	.821						
9	.034		.608						
12	.080	.434	.610	.983	.724				
24	.267	.360	.437	.575	.732		.886		
36	.557	.496	.528	.630	.767	.971	.831		
48	.700	.536	.514	.593	.675		.935	.550	
60	.716	.549	.510	.555	.592		.691		
120	.008	.009	.010	.012	.011		.011	.007	.000

Table 9: Probability Estimates of λ Came From the Hypothetical Data									
n\m	1	2	3	4	6	9	12	24	60
2	.326								
3	.004								
4	.000	.000							
6	.837	.067	.196						
9	.567		.730						
12	.590	.865	.489	.524	.876				
24	.768	.865	.490	.434	.529		.014		
36	.927	.885	.634	.629	.939	.441	.019		
48	.935	.902	.613	.618	.936		.033	.071	
60	.951	.910	.625	.645	.925		.048		
120	0.972	0.924	0.797	0.827	0.931		0.205	0.308	0.000

Figure 1: The U-shape or Smile

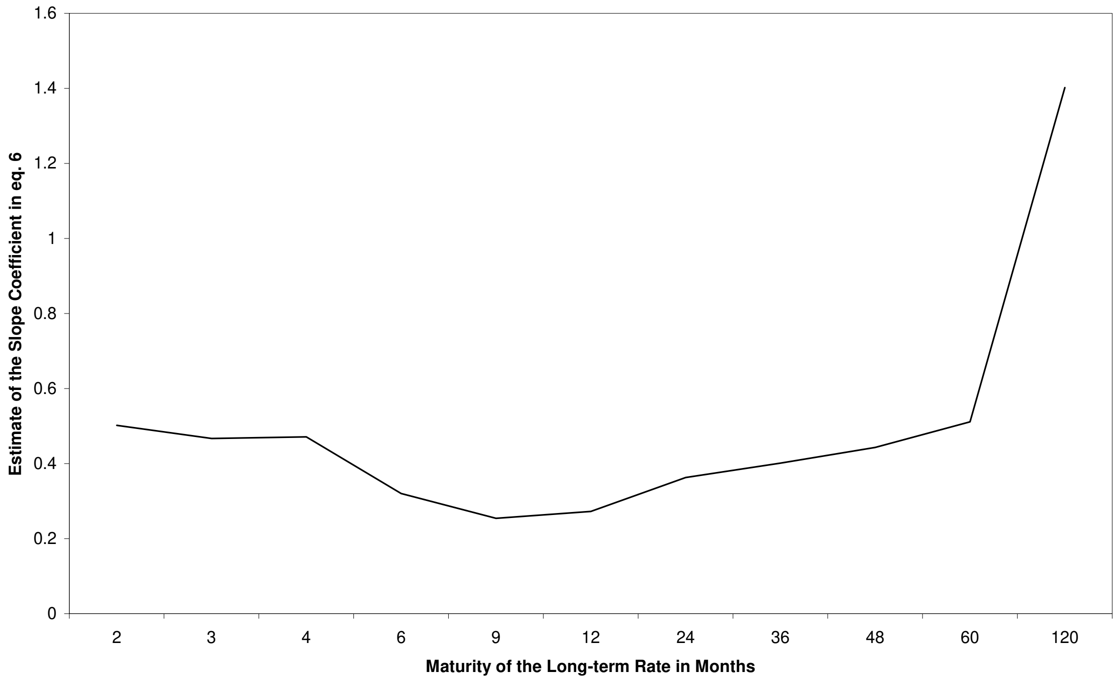

Figure 2: The Smirk

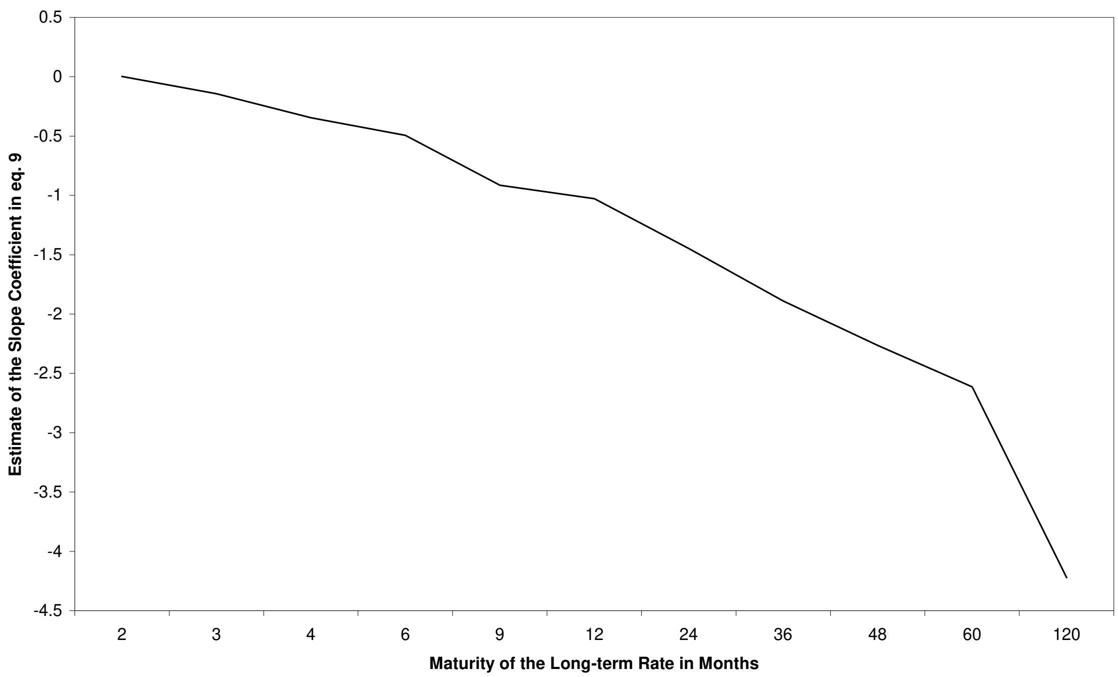

Figure 3: The Smile

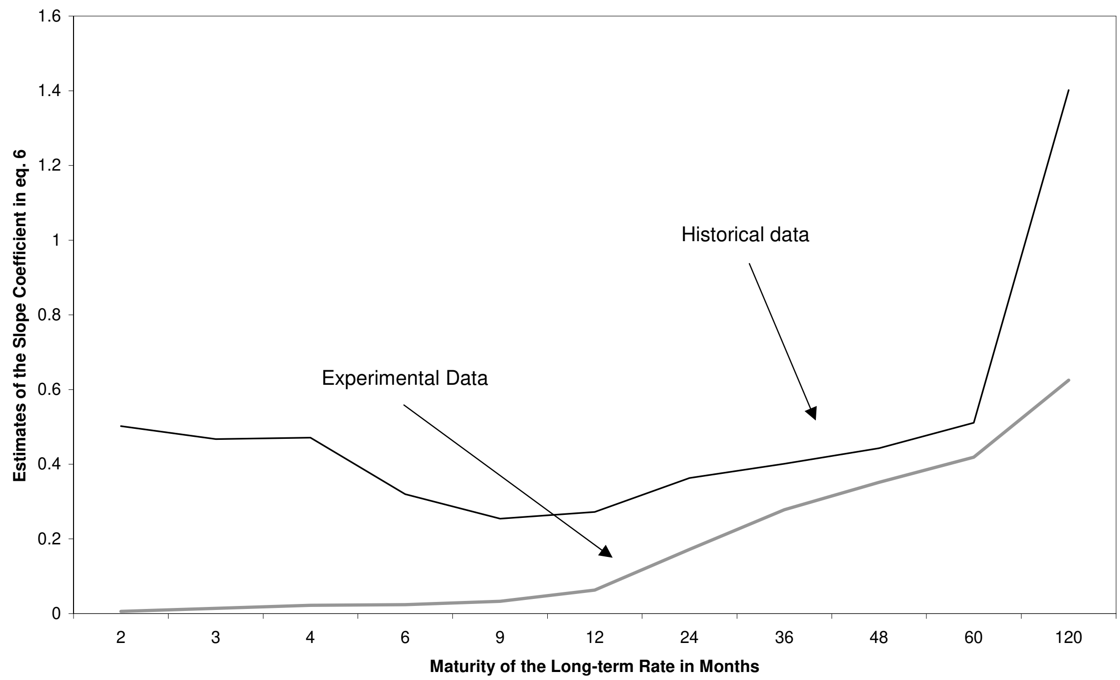

Figure 4: The Smirk